

Chapter 1

Referential cohesion in Bunun: A comparison of two genres

Rik De Busser

National Chengchi University

This chapter investigates how referential expressions are involved in establishing and maintaining textual cohesion in Bunun, an Austronesian language of Taiwan, and how this behaviour varies across genres. Relying on a model of referential cohesion inspired by systemic-functional grammar, it explores differences and similarities for encoding referential continuity across sentence boundaries in oral and narrative text. It concludes that, contrary to initial expectation, and despite considerable formal differences in how referential expressions are realized, at a more fundamental level the properties of referential cohesion are unexpectedly stable across genres.

1 Introduction

1.1 Cohesion

Now more than four decades ago, Halliday & Hasan (1976) published their seminal work on the linguistic subsystem that helps creating coherent text by establishing connections between related semantic elements in that text. More specifically, it is “a set of lexicogrammatical systems that have evolved specifically as a resource for making it possible to transcend the boundaries of the clause — that is, the domain of the highest-ranking grammatical unit” (Halliday & Matthiessen 2004: 532). They referred to this subsystem as cohesion and to the connections as cohesive ties, and described it as the set of “relations of meaning that exist within the text, and that define it as text” (Halliday & Hasan 1976: 4). This means that its realization is not confined by clause or other grammatical boundaries, but typically operates on the scale of text or discourse.

It also implies that there is no isomorphic relationship between grammatical devices and cohesive effects; cohesion pertains to semantic relationships within texts that “may take any one of various forms” (Halliday & Hasan 1976: 13). The original proposal, which has been integrated in Halliday’s systemic-functional grammar (Halliday 1994; Halliday & Matthiessen 2004; Halliday & Matthiessen 2014), distinguishes four types of relationships:



1. REFERENCE establishes cohesive ties between linguistic elements through various forms of spatio-temporal and personal deixis, and through comparison.
2. ELLIPSIS covers all phenomena that establish cohesive links by omitting a grammatical unit, or by swapping it for a placeholder element.¹
3. CONJUNCTION creates logical or spatio-temporal ties between propositions, typically through various grammatical mechanisms for clause linking.
4. LEXICAL COHESION is established between lexical elements through repetition and various semantic relationships.

The markers for each of the four types of cohesive relationship are indicated in the following examples.

(1) English

a. Reference

***That** man's dog is **much larger than** my cat.*

b. Ellipsis

*How many cookies are left? I took twelve Ø. **So did** you.*

c. Conjunction

***When** it shut down, something went wrong. **In short**, it caught fire.*

d. Lexical organization

Emperor penguins protect their chicks from the cold by putting the little fluff balls on their feet.

In (1a), *that* points to a referent that exists outside the text (exophoric reference), the phrase *much larger than* connects *that man's dog* and *my cat*, and the possessive form *my* creates an exophoric link to the speaker. In (1b), ellipsis in the second clause indicates that the head of *twelve Ø* refers to the same set of referents as *cookies* in the first clause. The substitutive construction *so did* in the third clause indicates that its subject performed the same action, *take [cookies]*, as the first person in the second clause. In (1c), *when* creates a relationship of simultaneity between the first and second clause, and *in short* indicates that the third clause summarizes the previous discourse. Finally, the penguin-related lexical items in (1d) arrange themselves in a complex of lexical cohesive relationships (see Figure 1).

¹Halliday & Hasan (1976: 88–141) called the latter substitution and originally considered it to be a separate cohesive category, meant to account for forms like *one* in English expressions such as *You can choose the blue candy or the red one*. They acknowledged that both substitution and ellipsis established cohesive ties by replacement, either by zero or by a placeholder (Halliday & Hasan 1976: 88) and both are subsumed under ellipsis in Halliday (1994) and later publications.

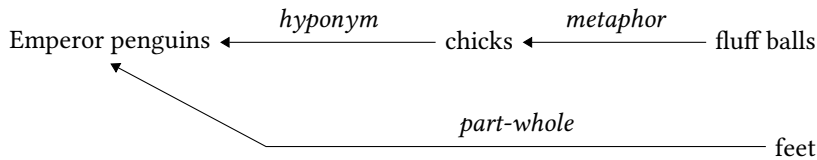


Figure 1: Schema of a cohesive chain

As the examples illustrate, markers of cohesion are highly heterogeneous in their grammatical and relational properties. What they have in common is that they establish cohesive ties, that is, semantic connections between linguistic elements (words, phrases, clauses, etc.) that are typically asymmetrical and express that the discourse segments in which they occur are to be interpreted as part of a coherent whole. These ties, either in isolation or by combining into longer chains, weave through a text. Together with thematic structure (theme/rheme contrasts) and focus structure (given/new), cohesion thus creates ‘texture’ (Halliday 1994: 334; Halliday & Matthiessen 2004: 579), the perception of a text or discourse as a connected whole. Texture in turn is “one aspect of the study of coherence, which can be thought of as the process whereby a reading position is naturalized by texts for listener/readers” (Martin 2001: 35).

Importantly, this implies that Hallidayan systemic-functional grammar “does *not* equate cohesion with coherence” (Martin 1992: 27; see also Martin 2001). Cohesion is merely one of the linguistic systems responsible for textual coherence. Later work on coherence often merged the two concepts, and typically reduced the phenomenon to a semantic-pragmatic component responsible for combining clause-level propositions into larger rhetorical structures (see e.g. Mann & Thompson 1987; Kehler 2002; Kehler 2004).

In sum, cohesion is an information-structuring device that, by establishing semantic connections between a heterogeneous set of linguistic units within a text, assists language users in interpreting that text as a cohesive, connected whole. In doing so, it is one of the subsystems responsible for structuring the distribution of information elements on a textual (supra-sentential) level.

Cohesion has been explored extensively in theoretical and applied linguistics, but overwhelmingly in the context of English (Halliday & Hasan 1976; Connor 1984; Martin 1992; Abadiano 1995; Tanskanen 2006; Crossley & McNamara 2012) and occasionally other major languages (Aziz 1988 on Arabic; Hickmann & Hendriks 1999 on English, French, German and Mandarin; Kruger 2000 on Afrikaans; Hassel 2005 on English, German and Norwegian). Work on minority languages is much less common. In the Austronesian world, the only studies I am aware of are Ezard (1978) on Tawala, Flaming (1983) on Wandamen, and Benn (1991) on Central Bontoc. The first two are literal applications of Halliday & Hasan’s framework to their languages; Benn employs a number of frameworks, including Halliday & Hasan’s, for his analysis of the discursive structure of Central Bontoc ritual texts.

This chapter adapts Halliday & Hasan's original model to fit the needs of analysing the role of referential expressions in establishing the cohesive texture of Bunun texts. It investigates the role cohesion plays in establishing genre distinctions through a small-scale pilot study.

1.2 Genre and cohesion

Genres or registers can be defined as specific types of texts or discourse with sets of "relatively stable" properties associated with the "thematic content, style, and compositional structure" that reflects the specific needs of well-defined contexts in which they were realized (Bakhtin 1986: 60).

Distinctions between genres are marked through various linguistic means. Biber (1995: 28) makes a basic distinction between register markers and register features. The former are linguistic cues that are specific to a certain register or genre and therefore directly indicate that a text belongs to it. An anecdotal example is the phrase *a long time ago in a land far, far away* introducing a fairy tale. Register features are linguistic elements that are not genre-specific, but whose frequency or distribution is in certain situations indicative of a specific register or genre. For instance, imperatives are relatively common in recipes, but they occur in many other genres as well. Cohesion falls into the latter category.

Research on the relationship between genres and their indicators has mainly focussed on the "relative distributions of surface linguistic features, such as adjectives, nominalizations, passives, and various clause types" (Biber 1995: 12). Even Biber, who went well beyond previous studies by focussing on complex feature bundles, mainly concentrates on morphosyntactic features that can be straightforwardly extracted from the surface realization of the text (see Biber 1995: 94–104; also Biber & Conrad 2009: 217–226). Given that genre is associated with the global discursive and semantic features of texts, one should probably assume that these grammatical features serve as proxy indicators of certain structural elements of meaning, discourse organisation and information structure.

Cohesion is an important determinant of the distribution of information in text, so it is reasonable to assume that it is interconnected with the global properties of text structure, and therefore contributes to (Halliday & Hasan 1976) or closely interacts with genre (Martin 1992; Martin 2001). There are a number of reasons why one would expect consistent correlations between cohesion and genre, many linked to the accessibility of linguistic information (Lambrecht 1994: 74–116; Ariel 1991).

First, expectation patterns related to the nature and quantity of assumed background knowledge and explicitly expressed information are often genre-specific. Specialized genres assume a greater volume of background knowledge than more generalized genres. For example, the presupposed background knowledge in an informal conversation is different from that in an academic textbook (Biber & Conrad 2009: 14–15). This affects the need for explicitly expressing cohesive relationships between elements in a text.

Second, differences in genre commonly correlate with differences in modes of realization, which in turn influences the options for realizing cohesion. For instance, oral and written genres diverge in which cohesive strategies they employ (see e.g. Fox 1987; Givón 1993). Textual coherence in writing is partly realized through meta-linguistic means, such as writing conventions and punctuation, not available in oral discourse. Because of the visual nature of the written medium, information is also more readily, and longer, accessible. All things being equal, one would therefore expect that oral genres tend to have a more dense cohesive structure (or a larger presence of other coherence-creating mechanisms) than their written equivalents, in order to reach an equal level of coherence.

For certain types of cohesive relationships, the link between cohesion and genre is well understood. For instance, it is uncontroversial that “genre-specific conventions [...] play a significant role in anaphoric patterning in conversation and writing” (Fox 1987: 2). Research explicitly comparing cohesive patterning across genres is scarce, but the influence of cohesion on the realization of individual genres is the subject of a number of studies. The above-mentioned Benn (1991) investigates cohesion in single genre (written essays) in Central Bontoc. Another example are Malah & Rashid (2016) who, based on Hoey (1991), explore the role of content words in establishing the cohesive properties of English language Nigerian newspaper texts.

One important question is in which manner exactly cohesion indicates genre in text or – from a comparative perspective – how its realizations are indicative of differences between genres. Halliday & Hasan suggest that “the possible differences among different genres and different authors [are] in the numbers and kinds of tie they typically employ” (Halliday & Hasan 1976: 4). In other words, one would expect that (1) cohesive density, and (2) the nature of the connections between elements in cohesive relationships varies between genres.

This chapter investigates whether, and to what extent, these two hypotheses are true for two text genres, oral narrative and biblical translations, in a Bunun speech community. It compares the cohesive density and the morphosyntactic and semantic-pragmatic properties of cohesive ties in these two types of texts.

Similar to Malah & Rashid (2016), this chapter does not discuss all aspects of cohesion as they were introduced in Halliday & Hasan (1976). Rather, the discussion focuses on referential cohesion, the conceptually coherent subset of cohesive ties that is involved in establishing relationships between referential items. Its exact delineation is discussed in §1.4. Before this is possible, I first introduce the Bunun language, its dialects and the genres involved in the present analysis.

1.3 Bunun dialects

Bunun is one of around sixteen Austronesian languages spoken on Taiwan (Li 2008). It has five extant dialects that are classified into a Southern (Isbukun), Central (Takbanuaz and Takivatan) and Northern branch (Takibakha and Takituduh). Within the Isbukun dialect, at least three distinct varieties are spoken in Kaohsiung, Taitung, and Nantou. Be-

tween dialects, especially between Isbukun varieties and dialects of the other branches, there is a fair amount of phonological, lexical, and grammatical differentiation (see Li 1988 for an overview of phonological and lexical variation). Only the Takivatan and Isbukun dialects are relevant to the present discussion.

(2) Bunun

- a. Takivatan (fieldwork, observed)
mun-ʔisaq=ʔas
 ALL-where-2S.SUBJ
 ‘Where are you going?’
- b. Isbukun (Lilian Li, pers. comm.)
ku-ʔisaʔ kasuʔ ma-tuktuk lukis
 ALL-where 2S.SUBJ DYN-chop wood
 ‘Where do/did you go to chop wood?’

Example (2a–2b) illustrates the degree of discrepancy between the two dialects.² The coda of the question word (/q/ in Takivatan, /ʔ/ in Isbukun) is illustrative of a systematic phonological contrast. In near-identical contexts, both dialects use different allative prefixes. Finally, whereas Takivatan prefers a pronominal clitic in subject positions like this, Isbukun uses a free pronoun that does not exist in Takivatan (see Table 1 and Table 2).

Bunun dialects have a verb-initial constituent order and what has been called a Western Austronesian or Philippine-type voice system (see French 1987; Foley 2007; Riesberg 2014 for general overviews), which in Bunun distinguishes at minimum between actor (AV), undergoer (UV), and locative voice (LV), marked by suffixes on the verb. In (3a), *siða* is actor voice and as a result unmarked; the UV in (3b) is indicated by a suffix *-un*, and the LV in (3c) by *-an*.

(3) Takivatan Bunun

- a. Actor voice ((fieldwork, elicited))
na-siða qaimaŋsuð.
 IRR-take thing
 ‘I will pick up things.’
- b. Undergoer voice ((fieldwork, text corpus))
maŋmaŋ ni siða-un.
 many NEG take-UV
 ‘... a lot were not caught.’
- c. Locative voice ((fieldwork, text corpus))
maqtu pa-siða-an-in ŋabul vanis.
 can CAUS.DYN-take-LV-PRV deer wild.boar
 ‘... and we could catch deer and wild boar.’

²The following changes were made to graphemic conventions: z > ð, ’ > ʔ, ch > ɕ, ng > ŋ

Certain analyses additionally include instrumental, beneficiary, and resultative object voices, but these forms are relatively uncommon and can be further ignored here.

The remainder of this section gives a short overview of various deictic paradigms, since these are relevant to the discussion at hand. All five Bunun dialects have sets of free and bound personal pronouns. Paradigmatic distinctions are largely equivalent, but the pronominal sets have formally diverged and have been analysed as expressing different grammatical distinctions in Takivatan and Isbukun. Tables 1 and 2 give the pronominal paradigms for both dialects.³

Table 1: Personal pronouns in Takivatan Bunun

(a)	Subject		Non-subj.		Poss.	Loc.
	Free	Bound	Free	Bound		
1S	sak, saikin	-(ʔ)ak	ðaku, nak	-(ʔ)uk	inak, ainak, nak	ðakuʔan
2S	—	-(ʔ)as	suʔu, su	—	isu, su	suʔuʔan
3S	(see b)	-(ʔ)is	(see b)	—		
1I	ʔata, inʔata	—	mita	—	imita	mitaʔan
1E	ðamu, sam	-(ʔ)am	ðami, nam	—	inam, nam	ðamiʔan
2P	amu	-(ʔ)am	muʔu, mu	—	imu, mu	muʔuʔan
3P	(see b)	—	(see b)	—		
(b)	Subject & Non-subject					
	PROX		MED		DIST	
3S	isti		istun		ista	
3P	inti		intun		inta	

Some of the more systematic differences are worth mentioning. Third person pronouns in Takivatan differentiate between proximal, medial and distal forms and do not have distinct subject and non-subject forms. They can therefore be interpreted as a subset of demonstratives (Table 3). In contrast, Isbukun third person pronouns do not encode a deictic contrast. Singular forms all appear derived from the stem *sia*, which in Takivatan is an anaphoric form that appears in a number of grammatical positions (De Busser 2009: 467–474). Plural forms all derive from the stem *nai*. Zeitoun (2000: 72) suggests

³The Takivatan data on personal pronouns is from De Busser (2009: 441); the Isbukun data from the Kaohsiung variety in Huang & Shih (2016: 85). The latter mark vowel length by grapheme doubling. This distinction is non-phonemic in Bunun: generally, monosyllabic roots tend to have lengthened vowels, irrespective of the environment in which they occur. To make comparison easier, long vowels in the Isbukun examples are represented by single vowel graphemes. Subject and non-subject forms are analysed and glossed differently in De Busser (2009) and Huang & Shih (2016), in this might reflect subtle differences in the grammatical distribution of these forms. Again, to make comparison easier, this terminology has here been homogenized.

Table 2: Personal pronouns in Isbukun Bunun

	Subject		Non-subj.		Poss.	Loc.
	Free	Bound	Free	Bound		
1S	saikin	-ik	ðaku	-ku	inak	ðakuan
2S	kasu(n)	-as	su	-su	isu	suan
3S	saia, sai(n)	—	saidɕia	—	isaidɕia, isia	siʔaan ɕia
1I	kata	-ta	ita, mita	-ta	imita	mitaan
1E	kaimin	-im	ðami	—	inam	ðamian
2P	kamu(n)	-am	mu	-mu	imu	muan
3P	nai, nian	—	nai (VIS),	—	inai (VIS),	naian ɕia
	(VIS), naia (NVIS)		naiɕia (NVIS)		inaidɕia (NVIS)	

that variant forms within each category code a visibility distinction. The element *ɕia* on third person forms is in all likelihood a distal determiner enclitic, making their status of as personal pronouns contentious.

Demonstrative pronouns vary substantially between dialects. De Busser (2017: 95–97) describes an elaborate paradigm for Takivatan; see Table 3.

Table 3: Free demonstratives in Takivatan Bunun

		PROX	MED	DIST	USPEC
S	VIS	aipi	aipun	aipa	aip
	NVIS	naipi	naipun	naipa	naip
P	VIS	aiŋki	aiŋkun	aiŋka	—
	NVIS	naiŋki	naiŋkun	naiŋka	—
GNR	VIS	aiti	aitun	aita	—
	NVIS	naiti	naitun	naita	—
PAUC	VIS	—	—	(ainta)	—
	NVIS	—	naintun	(nainta)	—

None of these forms has so far been attested in Isbukun. The demonstrative forms described in Huang & Shih (2016: 95) distinguish case and distance, but not visibility. However, their paradigm consists of fully transparent combinations of the form *sia* or the spatial adverbs *di* and *adi* ‘there’ with various bound determiners (see Table 4), which encode both distinctions mentioned above.

Finally, all Bunun dialects have two sets of bound determiners, which encode a distance contrast and can occur on a variety of word classes including verbs (see De Busser 2009: 427–440 for an explanation).

Table 4: Determiners in Isbukun and Takivatan Bunun

		Takivatan	Isbukun (Taitung & Kaohsiung)	Isbukun (Nantou)
D1	PROX	-ki	-in	-in
	MED	-kun	-an	-an
	DIST	-ka	-a	-a
D2	PROX	-ti	-ɕin	-tin
	MED	-tun	-tan	-tan
	DIST	-ta	-ɕia	-tia

Again, there are formal differences, this time even between different varieties of the Isbukun dialect. The distinction between D1 and D2 appears to be fundamentally different in the two dialects. De Busser (2009: 426–440); De Busser (2017) argues that in Takivatan the difference between the two sets is semantic in nature; in Isbukun, the difference is associated with case (D1 = NOM, D2 = OBL; see Huang & Shih 2016: 95; Zeitoun 2000: 76). Bound determiners are optional, and are considerably more common in Takivatan than in other dialects.

The paradigms above serve as illustrations of the degree of differentiation between Takivatan and Isbukun, and give an overview of some of the paradigms that are relevant to the present analysis, as they directly influence the difference between different text genres in Bunun.

1.4 Narrative genres in Bunun

This chapter compares two narrative genres, traditional oral narrative text and biblical narrative, as they occur in a single Takivatan Bunun language community in the village of Bahuan at the East Coast of Taiwan.

Traditional narrative texts, despite being transmitted orally, are by nature not improvised. In traditional Bunun communities, both expository and narrative texts in formal settings follow relatively strict conventions that govern amongst other things: (1) who has the right to speak and when; (2) how certain traditional knowledge should be presented; and (3) which formal aspects, such as formulae related to politeness and the veracity of the narration, should be included in specific oral genres. Many of these conventionalized aspects of stories appear to be the result of an ongoing consultation process between the elders, or a larger group of members, of the community. Transgressions of these rules are usually pointed out by authoritative members of the community, usually male elders.

A second genre with which many Bunun people are confronted on a nearly daily basis are biblical narratives. Presbyterian and Catholic missionaries introduced Christianity after the end of the Second World War, and it is an important part of contemporary Bunun culture. From the 1940s onwards, especially the Presbyterian Church, through

the Bible Society of Taiwan, has been active in translating biblical texts into the Bunun language.

Bible translations are intricate undertakings that typically involve a translation team consisting of translators, native language consultants, and theologians. Especially in cultures that have little historical affinity with the Judaeo-Christian tradition, this process is more than simply translating texts: it requires the meticulous transposition of an alien conceptual universe with its associated lexical and grammatical framework (see De Busser 2013). This makes it nearly impossible to produce translations for every dialect of Bunun.

The present Bunun Bible translation (Bible Society in Taiwan 2000) will be referred to as the Bunun Bible. It is the authoritative translation containing the full New Testament and an abbreviated Old Testament, and is heavily based on the Isbukun dialect. Despite this, it is used in almost all Bunun churches, irrespective of their denomination or the dialect area to which they belong. From a language planning perspective, this made a lot of sense: Isbukun is the largest dialect, has the widest geographical spread, and has been studied most extensively. This is not to say that the Bunun Bible is a written mirror of any specific Isbukun variety: the translation process rather resulted in a supra-dialectal written standard for Christian texts in Bunun, which is also used for other religious text genres such as hymns.

An interesting consequence is that, although many Bunun are reading or listening to excerpts of the Bible on a regular basis, especially “to members of Takbanuaz, Takivatan, Takibakha and Takituduh communities, the language of their Bible is not closely related to the common vernacular” (De Busser 2013: 67). Since the discrepancy between Isbukun and the Takivatan dialect is considerable, the result is a situation in which two dominant narrative genres in the Takivatan language community have relatively divergent dialectal characteristics.

This leads to a question: how and to what extent do language users in non-Isbukun communities interpret the content of these texts that belong to related genres but have quite distinct lexical and grammatical characteristics? To an extent, this is a matter of lexical and grammatical overlap between dialects. However, an additional question concerns the cohesive fabric of these two types of texts. Given the differences between grammatical paradigms that are central to establishing cohesion, such as personal pronouns, demonstratives, and determiners, how do language users keep tab of issues such as thematic integrity, topic continuity, and the general distribution of information in a narrative progression?

The present study will investigate this issue by looking at how these two different narrative genres realize cohesion, and in a particular subset of cohesive relations that is here referred to as referential cohesion.

1.5 Referential cohesion

The basic framework for my analysis of referential cohesion is set out in De Busser (2017). Halliday & Hasan (1976) consider cohesion to be the aggregate set of cohesive ties, se-

mantic relationships that exist between meaningful elements in a text. Cohesive ties are directional: they point from a textual source, which will be called the Reference (Rc), to a second element, which can exist inside (4a) or outside the text (4b).

- (4) a. Rc points to text-external Rt
maq aipa Rc [\rightarrow Rt: object in external reality]
 what DEM.S.DIST.VIS
 ‘What is that?’
- b. Rc points to text-internal Ta
maq-a ainak-a tama-ka tu-tuđa tu miqđiq daiŋ?ađ
 DEFIN-LNK IS.POSS-LNK father-K.DIST INTENS-real COMPL difficult large
aipa. Ta \leftarrow Rc
 DEM.S.DIST.VIS
 ‘As for my father, he really had a lot of difficulties.’

When this element is linguistic in nature, I call it a Target (Ta). Targets are themselves References that point back to previous targets. In doing so, they create cohesive chains, referential strands of different length that ‘weave’ through a text. Eventually, the final Reference of each chain points to a concept that exists outside the textual universe; this is called the Referent (Rt) of the cohesive chain and is in effect its ultimate Target (Halliday & Hasan 1976: 329; De Busser 2017: 107–108). This is schematised in Figure 2.

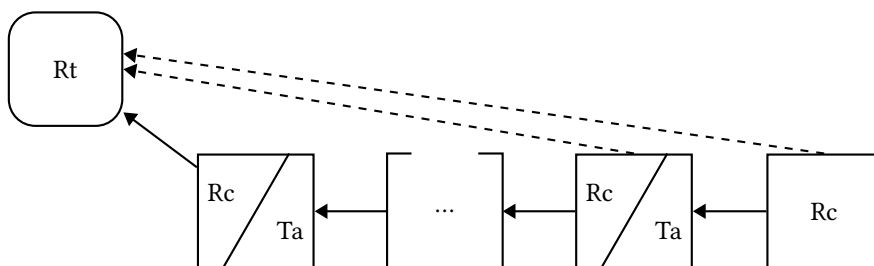


Figure 2: Schema of a cohesive chain

Together with other linguistic mechanisms, such as prosody, event expression, and contextual information, cohesion allows language users to interpret a text as a coherent whole operating in a context. It does this by expressing “the continuity that exists between one part of the text and another” (Halliday & Hasan 1976: 299). In doing so, cohesion forms an interface between the local and global distribution of information elements in a text. This interaction is given to be complex, but one specific example in which referential cohesion interacts with clause-internal information structural devices is by the “Theme tell[ing] the hearer where to start from in the interpretation of a message, and the cohesive ties signal how the message latches on to other parts of the discourse” (Hasselgård 2004: 68).

A crucial aspect of Halliday & Hasan's (1976) model is that cohesive ties are inherently semantically motivated, and therefore do not correspond to a single grammatical mechanism or exist between fixed classes of words or other linguistic elements. Somewhat contradictorily, Halliday & Hasan partly rely on grammatical criteria to distinguish different types of cohesion: they make a basic distinction between grammatical cohesion, which is expressed by grammatical means such as function words and grammatical constructions, and lexical cohesion, which is expressed through content words. These two categories further break down in an assortment of subtypes, based on a combination of semantic and grammatical criteria (see Halliday & Hasan 1976: 324).

This chapter takes a different approach; it focuses exclusively on referential cohesion, "the set of cohesive relations that create referring relationships between linguistic forms and referents" (De Busser 2017: 107). These are all linguistic expressions that can be targeted by deixis (or simply, that can be pointed at). In essence, this combines Halliday & Hasan's category of reference, with the exclusion of comparative reference, and their category of lexical cohesion, with the exclusion of collocation.

The criterion for establishing referential cohesion is semantic: all words and sub-lexical elements that are involved in establishing referential meaning are included in the cohesion analysis irrespective of their word class or grammatical status. Their involvement in reference is determined by their ability to be indicated by deictic expressions. The only formal restriction is that (with the exception of lexical compounds) phrases and other multi-word units are excluded, to prevent the data selection process from becoming too arbitrary. Word classes that have so far been implicated in referential cohesion in Bunun are: (1) nouns; (2) personal pronouns (see Table 1 and Table 2); (3) demonstrative pronouns (see Table 3); (4) bound demonstrative articles (see Table 4); (5) the anaphoric marker *sia* and its derivations; (6) numerals; (7) words expressing time, manner and location; (8) question words; and (9) certain verbal roots.

It is appealing to equate referential cohesion to anaphora resolution, but this is only so in a very broad sense. Phoric reference is typically seen as a grammatical property of language that is involved in referent tracking and uses fixed morphosyntactic strategies to establish relationships of identity between expressions and their antecedents (see for instance Huang 2000). On the other hand, referential cohesion, though obviously involved in reference tracking, is a semantic property of language that creates meaning relations between two referential expressions. These are not always relations of identity (although they can be), and neither do they necessarily have straightforward morphosyntactic correlates. For instance, the cohesive tie between *bantas* 'legs' and the preceding word *bunun* 'man' is meronymic in nature, and the fact that both are nouns is not grammatically determined.

(5) T3.8a

itu bunun-ɖia bantas mas buhtuŋ hai ɖinsu
 this.here people-DIST.OBL leg OBL joint TPC immediately
is-tamasað-an
 TRANSFER-strength-LV

'This man here, his legs and joints immediately became powerful, ...'

The fact that cohesion is involved in shaping the general structure of a text suggests that it varies between and is therefore indicative of genre (Martin 2001). One of the goals of this research is to try to establish the nature of this variation. One possibility suggested by Halliday & Hasan (1976: 4) is that “among different genres and different authors in the numbers and kinds of tie they typically employ.” Another option is that genres vary in cohesive density, the number of cohesive elements or cohesive chains relative to text length. I will investigate both possibilities in the next section by analysing three short text segments.

2 Cohesion analysis

2.1 Data

The present study is a small-scale comparison of oral and biblical narrative as it occurs in the Takivatan Bunun speech community in the village of Bahuan (Chinese name: Mayuan) at the East Coast of Taiwan. It consists of the analysis of three text excerpts (T1, T2, and T3), which are given in their entirety with cohesive elements underlined in the Appendices. All three are part of larger narratives; segment boundaries were chosen to extract internally coherent sub-narratives.

The first two texts, T1 and T2, are traditional oral narrative sequences. They feature two elderly men, both fluent speakers of Takivatan Bunun and both around 75 years old at the time of recording. Text T1 is an account of a hunt and is part of a long story in which the speaker, Vau Taisnunan, recounts his life story. In text T2, Tulbus Manququ recounts how the traditional Bunun hunters used prophetic dreams to determine the appropriate time for the hunt.

Text T3 is an excerpt from the Acts of the Apostles in the Bunun Bible (Bible Society in Taiwan 2000), in which the apostles Peter and Paul heal a cripple. As mentioned in §1.2, it is a written text that is the product of a complex translation process. The spelling of the original text in the Bunun Bible was adapted to make it consistent with the Takivatan Bunun texts and make it adhere to a one-grapheme-per-phoneme principle (see Footnote 1).

2.2 Methodology

All elements in these texts that could be unambiguously identified as having a referential function were marked for analysis. Importantly, this means that non-expressed (ellipted) elements are not included, despite having a referential value. In contrast with the coding scheme in Halliday & Hasan (1976: 329–355), no prior assumption is made about the word class (or morphological class) of the elements involved; so far only the nine linguistic classes mentioned in §1.5 have been attested in referential cohesive relationships.

In the data set the following information about each Reference, and the nature of its cohesive tie to its Target and Referent are encoded:

1. The location of a Reference in the text;

2. Its word class (see §1.5 and Table 5);
3. Its Target and the location of the Target in the text (this is not relevant for exophoric links);
4. The Referent of its cohesive chain, i.e. the text-external entity (or event) to which the Reference eventually refers;
5. The ontological type of the Target, i.e. whether the immediate Target of a Reference refers an event, a location or time, or a textual element (see Table 6);
6. The phoric status of the cohesive tie, i.e. whether it is an anaphoric, cataphoric, or exophoric link (see Table 7);
7. The relationship between the concept indicated by a Reference and the concept indicated by its Target (see Table 8).

The first four data points provide information about the structural properties of cohesive chains; the information in 5–7 relates to conceptual and informational properties of individual cohesive ties. The Referent (data point 4) of referential expressions is indicated by unique names that allow us to track cohesive chains. Targets of cohesive ties can belong to a number of distinct ontological types (data point 5). Most commonly they are concrete or abstract entities in the real world, but they can also be reified events, physical or temporal locations, or textual elements; this is discussed in §3.2. The phoric status (data point 6) indicates whether a Reference points to a Target that precedes it (anaphoric reference), follows it (cataphoric reference) or exists outside the textual universe (exophoric reference).

Finally, data point 7 encodes the conceptual relationships between References and their Targets. Possible values are adapted from the set of relations subsumed under Halliday & Hasan (1976: 277–282) category of reiteration, complemented by Peirce's fundamental semiotic relationships that exist between signifiers and semiotic objects (metaphor, metonym, symbol; see Merrell 2001). The following relationships are distinguished:

Identity: Relations in which the Reference refers to the same concept as its Target. This can be because it is a literal repetition, a personal or demonstrative deictic, a synonym, or a near-synonym.

Hyponym/hyperonym/co-hyponym: Relations dictated in terms of category membership. Hyponyms refer to other concepts of which they are a subclass; hyperonyms refer to concepts of which they are a superclass; co-hyponyms are terms that have the same immediate superclass.

Part/whole/co-part: Relations that defined in terms of meronymy (see example 4 above).

Antonym: Relations based on conceptual opposition.

Metaphor: Relations based on similarity, other than identity.

Metonym: Relations based on proximity or, more generally, contiguity.

Symbol: Relations based on conventional semantic connections that cannot be reduced to any of the previous six relations.

Originally envisaged to be applicable to lexical cohesion alone, these relations here extend to all referential cohesive ties. In the present sample, no metaphoric and symbolic relations have been attested. The referential cohesion analysis of the sample texts is included in the Appendix.

The next section compares the distribution of these data in the three text samples in order to investigate the following questions:

1. How similar or different are oral (T1 & T2) and biblical narrative (T3) in terms of referential cohesion?
2. How do systematic differences manifest themselves?
3. Given that biblical texts are strongly influenced by the Isbukun dialect, to what extent are differences the result of dialect differentiation and to what extent of genre differentiation?

On a more fundamental level, these questions provide an insight in how Takivatan speakers deal with the genre innovation that biblical narrative has brought to their literary repertoire.

§3.1 discusses the distribution of word classes in the three texts (data point 2), and §3.2 that of various functional properties (data points 5–7). In §3.3, I look at the global properties of cohesion in oral and biblical narrative text. Note again that this exploratory study uses a relatively small text sample.

3 Discussion

3.1 Word class

Let's first have a look at how different word classes are involved in the expression of referential cohesion. Table 5 shows the distribution of word classes of References in the three text excerpts used in the present analysis (please take into account that these results are indicative only).

In line with expectation, nouns are the dominant word class by a considerable margin in all texts and both text types: referential cohesion prototypically involves reference to concrete or abstract entities, and cross-linguistically these are typically expressed by nouns. However, the relative proportion of nouns is significantly higher in biblical text than it is in oral narratives (44.59% in T3 vs. 25.45% in T1+T2). In oral narratives, this relative scarcity of nouns is offset by a relative abundance of place and time words. In

Table 5: Word class of References (Rc) of cohesive ties

	Oral narrative						Biblical narr.	
	T1		T2		T1+T2		T3	
	%	#	%	#	%	#	%	#
anaph. marker	10.71	3	3.70	1	7.27	4	5.41	4
article	14.29	4	14.81	4	14.55	8	4.05	3
dem. pronoun	10.71	3	0	0	5.45	3	0	0
place word	10.71	3	14.81	4	12.73	7	2.70	2
manner word	0	0	0	0	0	0	0	0
noun	21.43	6	29.63	8	25.45	14	44.59	33
numeral	3.57	1	0	0	1.82	1	2.70	2
pers. pronoun	3.57	1	3.70	1	3.64	2	29.73	22
time word	14.29	4	7.41	2	10.91	6	1.35	1
question word	0	0	0	0	0	0	1.35	1
verb	10.71	3	25.93	7	18.18	10	8.11	6
TOTAL	100	28	100	27	100	55	100	74

the absence of any indication that these word classes behave differently in the Takivatan and the Isbukun dialect, the most likely explanation for these discrepancies is that it is a genre distinction. In oral narratives, especially when they concern historical accounts of a personally experienced past, the temporal and geographical anchoring of events is probably more important than in stories of a distant past that are mainly intended as moral lessons. Conversely, Biblical narrative often puts great emphasis on the symbolic significance of names and places; more than half of the nouns in T3 are proper names. This is evident when the distributions of ontological types of Targets are compared in both genres (Table 6 below).

A second categorial inversion between the two genres can be observed in the distribution of demonstrative (anaphoric markers, determiners, and demonstratives) and personal deixis (personal pronouns). In oral narratives, demonstrative reference accounts for 27.27% of all referential expressions, and person pronouns for a mere 3.64%. In the biblical excerpt, personal pronouns make up 29.73% of all References, and the three demonstrative classes combined only 9.46%. It is not clear how this discrepancy can be explained as a genre distinction: oral narratives represent a highly speaker-centric form of storytelling and one would assume a relatively high proportion of personal pronominal reference. In this case, the difference is more likely due to dialect variation. For instance, as suggested in §1.3, bound determiners are much (the data suggests three times) more common in Takivatan than in other dialects including Isbukun. We also saw that, whereas Takivatan has a highly developed free demonstrative paradigm, putative demonstratives in Isbukun are all complex forms involving deictic determiners. Conversely, the Isbukun

pronominal paradigm is more complex than that in Takivatan. In the sample, the most highly developed deictic paradigm also has the highest relative frequency in each language variety.

Interestingly, verbs sometimes express referential cohesion. This happens most commonly with verbs derived from nouns or locative words.

- (6) T2.1b
taŋus-aŋ mati-bahi.
 first PROG-have.prophetic.dream
 ‘... they interpreted a prophetic dream beforehand.’
- (7) T1.2b
mina-baʔav tupa naip tu:
 ABL-high.location say DEM.S.NVIS COMPL
 ‘... Tiang had returned, he had come back from the mountain and told us: ...’

3.2 Conceptual dimensions

This section investigates the distribution of referential cohesive elements in terms of their conceptual-semantic properties (data points 5–7 in §2.2). It first looks at the ontological type of the Target. As mentioned above, referential expressions indicate concepts that can be targeted by deixis. One would assume that the prototypical Target of a referential cohesive expression is a material entity of some sort. Table 6 shows that this is not always the case.

Table 6: Ontological type of the Target of the cohesive tie

	Oral narrative						Biblical narr.	
	T1		T2		T1+T2		T3	
	%	#	%	#	%	#	%	#
Entity	39.29	11	29.63	8	34.55	19	78.38	58
Event	–	–	18.52	5	9.09	5	1.35	1
Location / Time	60.71	17	48.15	13	54.55	30	16.22	12
Text	–	–	3.70	1	1.82	1	4.05	3
TOTAL	100	28	100	27	100	55	100	74

In biblical narrative (T3) entities indeed make up more than two-thirds of the Targets of cohesive reference, the majority unsurprisingly people or concrete objects in the narrative world, e.g. *Pitilu* ‘Peter’ (T3.1c), *naidʒia* ‘they < Peter and John’ (T3.3c) or *kim* ‘gold’ (T3.6b). However, in the oral narrative sample this is only one-third. More than half of Targets in T1 and T2 refer to a spatial or temporal location, such as *laqaiban* ‘route’ (T1.3b), *ʔita* ‘there (distal)’ (T2.2a) or *[tupa]-ka* ‘[say] at that time’ (T2.4g). In the previous

section, I already explained that this discrepancy is genre-related. Traditional oral narratives in Bunun culture are typically anchored in the immediate spatio-temporal context; in Biblical stories, on the other hand, identifying time and place is only of secondary importance relative to the need to keep track of people and objects that populate an unfamiliar narrative universe and commonly have a symbolic significance.

Counterintuitively, 9.09% of referential expressions in T1 and T2 and 1.35% in T3 refer to events. These generally are instances of event reification: events are reinterpreted as quantifiable objects, with a certain materiality and well-defined boundaries (Quine 1985; Zacks & Tversky 2001). Finally, a number of referential cohesive ties have a meta-textual function: rather than referring to anyone or anything in the narrative universe, they point towards part of the text itself. This type of cohesive tie corresponds to what Himmelmann (1996) and others refer to as discourse deixis. In Bunun dialects, these discourse deictic links are always expressed by *sia*, which in Takivatan, and possibly also in other dialects, is a specialized anaphoric marker and typically refers back to a phrase, clause or larger text segment in the immediate context (for shorter segments, typically the preceding sentence). For instance, in T3.3b the form *sia* in *sia masaniṣiṣiṣi pisvaṇḍuan* ‘the aforementioned Holy Temple’ refers back to an identical phrase in the previous sentence T3.2d, which in turn refers back to *masaniṣiṣiṣi pisvaṇḍuan* ‘the Holy Temple’ (T3.1c), the first mention of this particular Referent in the story.

Table 7: Phoric function of the cohesive tie

	Oral narrative						Biblical narr.	
	T1		T2		T1+T2		T3	
	%	#	%	#	%	#	%	#
Exophoric	10.71	3	25.93	7	18.18	10	16.22	12
Anaphoric	89.29	25	74.07	20	81.82	45	79.73	59
Cataphoric	–	–	–	–	–	–	4.05	3
TOTAL	100	28	100	27	100	55	100	74

Table 7 gives an overview of the distribution of phoric functions of the cohesive ties in the sample. Anaphoric reference is dominant in all genres: most referents central to the text are introduced near the beginning and tend to persist throughout the story. This also explains why exophora are less common: they often occur towards the front of the text. Cataphoric reference is rare and in the present sample is only attested in biblical narrative.

Finally, Table 8 gives a breakdown of the types of conceptual relationships between References and their Targets.⁴ It is important to realize that these relationships are conceptual rather than lexical semantic distinctions: they hold between the concepts indi-

⁴Totals in Table 8 do not add up to 100% because exophoric cohesive ties have no associated conceptual relationship.

cated by referential expressions, and not only lexemes, as is the case in Halliday & Hasan (1976). This makes it possible, for instance, to establish a part-whole relationship between the noun *?ima* ‘hand’ (T3.7a) and the pronoun *isaidzia* ‘3s.POSS’ (T3.7a).

Table 8: Cohesive relationship between Rc and Ta

	Oral narrative						Biblical narr.	
	T1		T2		T1+T2		T3	
	%	#	%	#	%	#	%	#
Identity	67.86	19	40.74	11	54.55	30	55.46	41
Hyponym	7.14	2	11.11	3	9.09	5	–	–
Hyperonym	–	–	7.41	2	3.64	2	–	–
Cohyponym	–	–	–	–	–	–	2.70	2
Part	–	–	3.70	1	1.82	1	8.11	6
Whole	3.57	1	–	–	1.82	1	2.70	2
Copart	–	–	–	–	–	–	1.35	1
Antonym	7.14	2	–	–	3.64	2	–	–
Metaphor	–	–	–	–	–	–	–	–
Metonym	7.14	2	7.40	2	7.27	4	9.46	7
Symbol	–	–	–	–	–	–	–	–
TOTAL	92.86	26	70.37	19	81.82	45	83.78	59

The introduction mentioned that referential cohesion is not necessarily identificational and is therefore not exclusively “concerned with resources for tracking participants in discourse” (Martin 2001: 38). However, from the data it is clear that this is an important aspect of cohesion: in both text genres, around 55% of all cohesive ties establish relationships of identity, and their function relates to reference tracking. Among the general semiotic relationships (metaphor, metonym, symbol), only metonymy is attested in the sample.

One possible minor difference between genres is that oral narrative appears to prefer hyponymic relationships, and biblical texts meronymy. However, this is in all likelihood an incidental difference resulting from the biblical story having as its main theme the miraculous healing of a physical handicap. Superficially, differences between oral and Biblical narratives appear almost non-existent, contrary to Halliday & Hasan’s (1976: 4) expectation that genres differ in the types of cohesive ties they employ.

In conclusion, despite differences between the frequency distribution of word classes in the two genres (see §3.1), and despite the fact that they have their origins in different dialects of Bunun, oral and biblical narratives are largely similar in terms of the relative distribution of phoric properties and types of cohesive ties. The most conspicuous difference between the two genres is in the ontological type of the Target: cohesive ties in oral narratives have a higher tendency to refer to location or time, biblical narrative tends to refer more to material entities.

3.3 Global properties

The final section of this discussion examines the global properties of referential cohesion in the three Bunun text samples. As mentioned, it has been asserted that one of the ways in which cohesion might exhibit genre-dependent variation is through consistent differences in its density. In other words, the “number and density of such networks is one of the factors which gives to any text its particular flavour or texture” (Halliday & Hasan 1976: 52). Biber (1995: 187–193) suggests that this is indeed the case for Korean: the degree to which cohesive relations, including referential cohesion, are explicitly expressed varies widely between text genres. In this study, I measure density in Bunun text in three different ways:

Referential density: The total number of words relative to the total number of References (or cohesive ties) in a text.⁵ Referential density gives a general impression of how much real estate cohesive referential expressions take up in a text. Note that it does not really measure which percentage of words are referential expressions, since References can be morphemes and a single word can therefore contain more than a single Reference (see e.g. *daiða-ki* ‘there-K.PROX’ in T1.2b).

Cohesive density: The number of cohesive chains in a text relative to the total number of words. This is a proxy indicator of what in the quote by Halliday & Hasan above is referred to as the density of the cohesive network, that is, how many cohesive chains weave themselves through a text of a normalized length.

Cohesive referential density: The number of cohesive chains in a text relative to the number of References. This measure indicates the average length of cohesive chains in a text, in terms of its average number of referential expressions.

Table 9 calculates these three density metrics for the three texts and the two genres in the present sample.

Table 9: Global properties of the text segments

	Oral			Biblical
	T1	T2	T1+T2	T3
# of words	62	80	142	179
# of referential expressions (Rc)	28	27	55	74
# of cohesive chains	7	9	16	19
Referential density (words / Rc)	2.214	2.963	2.582	2.419
Cohesive density (chains / words)	0.113	0.113	0.113	0.106
Referential cohesive density (Rc / chains)	4	3	3.438	3.895

⁵This measure is equivalent to Abadiano’s (1995: 308) cohesive density.

While referential density and referential cohesive density both seem to be vacillating around a central value, the most surprising result is that the value for cohesive density is almost completely equal (0.11) across texts and genres. Especially in a small sample, where a certain degree of instability is expected, it is not very likely that this is a spurious result. This is very much against initial expectation, as cohesive density is one of the factors that one would most expect to vary across text types. For instance, in planned written text, such as our biblical narrative, tracking entities and spatio-temporal locations is cognitively less demanding than in oral narration, where visual cues that allow the listener to reaffirm the status of activated concepts are not available. The basic assumption would therefore be that written narration does not need to be as cohesively dense as oral narration.

Not only is this not the case, the present sample suggests that cohesive density is a constant, at least in Bunun. This is the opposite of “the possibility of cross-linguistic universals governing the patterns of discourse variation across registers and text types” that Biber (1995: 359) is looking for: what we have here is a property of the supra-clausal information structure of language that appears to be impervious to personal or genre-based variation. The reasons for the stability of this value are at present unclear. One possibility is that languages have a tendency to evolve towards a cohesive equilibrium, in which texts are as cohesive as necessary to make them coherent but not more so, an equivalent on a textual level of Haiman’s (1983) competition between iconicity and expressiveness.

4 Conclusion

This leads us to an unexpected conclusion. Despite the evident grammatical differences between oral and biblical narratives in the sample, caused by dialect and genre differentiation, the conceptual properties of their underlying referential cohesive structures are surprisingly similar: against initial expectation, no major systematic differences can be observed in the phoricity or functional type of cohesive relationships. Even more so, the data suggests that, in defiance of lexical and grammatical variation in the two genres and dialects, the cohesive density of the two genres under investigation is invariant. This may point towards a cohesive constant underlying the structure of Bunun texts, though further research will need to verify this.

There are a small number of systematic differences between the two text genres. In terms of the referential type of the concepts they encode, referential cohesive ties in oral narrative tend to refer more to spatial or temporal location and those in Biblical narrative more to entities. This corresponds to a predilection for place and time words in the former genre, and for nouns in the latter. I argued above that this distinction is in all likelihood due to genre-based informational demands. On the other hand, a contrast in the frequency of demonstrative and personal deixis is probably rooted in dialect-related grammatical differences.

The present study is intended as a pilot, a fact-finding mission. Despite its modest data set, it has come up with interesting and unexpected results, but future research

is necessary to test whether the present results will stand when tested against larger, statistically valid and more diversified data sets, and to find out whether regularities can be found in any of the lower-level categories. A number of questions regarding the invariance in cohesive density need to be answered. Will the cohesive density constant hold up in a larger sample with more genre distinctions and dialects? If so, how can it be explained? And does a similar phenomenon exist in other languages?

Acknowledgments

The research in this chapter was made possible by grant 104-2410-H-004-139- from the Ministry of Science of Technology, Taiwan. Many thanks to the anonymous reviewers for their insightful comments and suggestions.

Abbreviations

?	function unknown	GNR	generic
1E	first person exclusive	HESIT	hesitation marker
1S	first person singular	HUM	human
2S	second person singular	INCH	inchoative
3P	third person plural	INTENS	intensifier
3S	third person singular	INTER	interjection
ABL	ablative prefix expressing movement from	IRR	irrealis
ALL	allative prefix expression movement toward	LNK	linker
ANAPH	anaphoric marker	LOC	locational prefix expressing position in or at
ART	article	LV	locative voice
ASSOC	associative	MED	medial
ATTR	attributive marker	N	noun
AV	actor voice	NEG	negator
CAUS	causative	NSUBJ	non-subject form
COMPL	complementizer	NUM	numeral
COORD	coordinator	NVIS	non-visible
CV	CV reduplication	OBL	oblique case marker
D1	determiner paradigm 1	P	plural
D2	determiner paradigm 2	PAUC	paucal
DEFIN	definitional marker	PLACE	place word
DEM	demonstrative	POSS	possessive
DIST	distal	PROG	progressive
DYN	dynamic	PRON	pronoun
EMO	emotive	PROX	proximal
ENUM	enumerator	PRT	particle
		PRV	perfective

Q	question word	TIME	time word
RECIP	reciprocal	TPC	topicalizer
RESOBJ	resultative object	USPEC	underspecified
S	singular	UV	undergoer voice
STAT	stative	V	verb
SUBJ	subject form	VIS	visual
SUBORD	subordinator		

Appendix

T1: Segment oral narrative

Source: Takivatan Bunun Corpus

Corpus location: TVN-008-002:130-134

Speaker: Vau Taisnunan, M, 75 y

Location and time: Bahuan (Mayuan), 2006

The excerpt below was previously published as example 22 in De Busser (2017).

Text

- (8) *Aupa tuḏa niaŋ tu nanu sanavan minsumina Tiaŋ, minabaʔav tupa naip tu:*
a. *aupa tuḏa ni-aŋ tu nanu sanavan min-suma-in-a Tiaŋ*
thus real NEG-PROG COMPL really evening INCH-return-PRV-LNK T.
b. *mina-baʔav tupa naip tu*
ABL-high.location say DEM.S.NVIS COMPL
‘But, when it wasn’t really evening yet, Tiang had returned, he had come back from the mountain and told us: ...’
- (9) *Na, maqtu laqbiŋina, naʔasa dusa ta matiskun, maluʔumi han baʔav daiḏaki, pinkaunun isian baʔavta, ŋabul.*
a. *na maqtu laqbiŋin-a na-asa dusa-ta ma-tiskun*
well be.possible tomorrow-LNK IRR-have.to two-T.DIST DYN-in.a.group
b. *maluʔum-i han baʔav daiḏa-ki*
disperse-PRT be.at high.location there-K.PROX
c. *pinkaun-un i-sia-an baʔav-ta ŋabul*
go.up-UV LOC-ANAPH-LV high.location-T.DIST deer
‘Well, tomorrow is possible, two of us will have to go together, and disperse when we get to this place, and we will climb upwards to the deer that is in that place above.’
- (10) *A, namaqaisaq dauka, saqnutai du sia ʔukai laqaiban.*
a. *a na-ma-qaisaq dau-ka*
INTER IRR-DYN-in.that.direction EMO-K.DIST

- b. *saqnut-ai-du sia ?uka-i laqaiban*
 get.stuck-PRT-EMO ANAPH NEG.have-PRT route
 ‘A, if he will go in that direction, he will get stuck there, without a way out.’
- (11) *Mei, mei kahaul duna ?uka duduma laqaiban, aupu tuḏa, maupa tupina.*
 a. *mei mei ka-haul dun-a*
 already already come.from.below line-LNK
 b. *?uka du-duma laqaiban*
 NEG.exist RED-other route
 c. *aupu tuḏa maupa tupa-in-a*
 thus real thus say-PRV-LNK
 ‘The track is coming from below, and there is no other way out, it really is like that, thus he told us.’
- (12) *Ansaisaṇa Atul Daiṇ tu “nis, matiṇmutin tamudana madav.”*
 a. *ansais-aṇ-a Atul daiṇ tu*
 forbid-PROG-ENUM A. large COMPL
 b. *ni-is ma-tiṇmut-in ta-mu-dan-a maḏʔav*
 NEG-3S.F STAT-morning-PRV ?-ALL-road-LNK embarrassed
 ‘But Big Atul forbade us: “no, when it has become morning, we will leave, it is embarrassing.”
- (13) *Na, ?ukin aipa ?ita namudanin, musbai naipa maqmut.*
 a. *na ?uka-in aipa ?ita na-mu-dan-in*
 well NEG.have-PRV DEM.S.DIST.VIS there.DIST IRR-ALL-go-PRV
 b. *mu-isbai naipa maqmut*
 ALL-cause.to.move DEM.S.DIST.NVIS night.time
 ‘Well, it will not be there anymore, it will be gone, it will have run away during the night.’

Cohesion analysis

Table 10 contains an analysis of referential cohesive elements in text T1 above. Numbers in the headers refer to the data points referred to in §2.2.

Table 10: Referential cohesion analysis T1

(1)	Reference (Re)	Re word class (2)	(3)	Target (Ta)	Referent (4)	Functional role Ta (5)	Phoric function (6)	Rel. Re-Ta (7)
1a	<i>sanavan</i> 'evening'	TIME			time of day	Loc/Time	Exophoric	
1a	<i>Tiaŋ</i> 'T'	N		(prev. text)	Tiaŋ	Entity	Anaphoric	Identity
1b	<i>[mina-]baʔav</i> 'come from the mountain'	V		(prev. text)	location deer	Loc/Time	Anaphoric	Identity
1b	<i>naip</i> 'DEM.S.NVIS'	DEM	1a	<i>Tiaŋ</i> 'T'	Tiaŋ	Entity	Anaphoric	Identity
2a	<i>laqbiŋin[-a]</i> 'tomorrow'	TIME	1a	<i>sanavan</i> 'evening'	time of day	Loc/Time	Anaphoric	Metonym
2a	<i>dusa-ta</i> 'two'	NUM		(prev. text)	we	Entity	Exophoric	Hyponym
2a	<i>[dusa]-ta</i> 'ART.ENT.DIS'	ART	1b	<i>[mina-]baʔav</i> 'come from the mountain'	location deer	Loc/Time	Anaphoric	Identity
2b	<i>baʔav</i> 'high location'	V	2a	<i>[dusa]-ta</i> 'ART.ENT.DIS'	location deer	Loc/Time	Anaphoric	Identity
2b	<i>daida[-ki]</i> 'that place'	PLACE	2b	<i>baʔav</i> 'high location'	location deer	Event	Anaphoric	Identity
2b	<i>[daida]-ki</i> 'ART.EVT.PROX'	ART	2a	<i>laqbiŋin</i> 'tomorrow'	time of day	Loc/Time	Anaphoric	Identity
2c	<i>i-sia-an</i> 'the place of that one'	ANAPH	2b	<i>daida-ki</i> 'that place'	location deer	Loc/Time	Anaphoric	Identity
2c	<i>[i-]sia[-an]</i> 'ANAPH'	ANAPH		(<i>dapana</i> 'foot prints') [008-002:125]	deer	Entity	Anaphoric	Whole
2c	<i>baʔav[-ta]</i> 'high location'	PLACE	2c	<i>i-sia-an</i> 'the place of that one'	location deer	Loc/Time	Anaphoric	Identity
2c	<i>[baʔav]-ta</i> 'ART.ENT.DIST'	ART	2c	<i>sia</i> 'ANAPH'	location deer	Entity	Anaphoric	Identity
2c	<i>ŋabul</i> 'deer'	N	2c	<i>-ta</i> 'ART'	deer	Entity	Anaphoric	Identity
3a	<i>[na-ma-qaisaq-dau]-ka</i> 'ART.EVT.DIST'	ART	2c	<i>baʔav-ta</i> 'high location'	location deer	Loc/Time	Anaphoric	Identity
3b	<i>sia</i> 'ANAPH'	ANAPH		<i>ŋabul</i> 'deer'	deer	Entity	Anaphoric	Identity
3b	<i>laqaiban</i> 'route'	N			route deer	Loc/Time	Exophoric	
4a	<i>[ka-]haul</i> 'below'	V	3a	<i>[na-ma-qaisaq-dau]-ka</i> 'ART.EVT.DIST'	location deer	Loc/Time	Anaphoric	Antonym
4a	dun 'line'	N	3b	<i>laqaiban</i> 'route'	route deer	Loc/Time	Anaphoric	Identity
4b	<i>laqaiban</i> 'route'	N	4a	dun 'line'	route deer	Loc/Time	Anaphoric	Identity
5a	<i>Atul daiŋ</i> 'Big Atul'	N		(<i>nas-Atul daiŋ</i> 'the erstwhile Big Atul') [008-002:126]	Atul	Entity	Anaphoric	Identity
5b	<i>[ni]-is</i> '3s.F'	PRON	3b	<i>sia</i> 'anaph'	deer	Entity	Anaphoric	Identity
5b	<i>[na-]tiŋmuŋ[-in]</i> 'morning'	TIME	2b	<i>[daida]-ki</i> 'ART.SIT.PROX'	time of day	Loc/Time	Anaphoric	Hyponym
6a	<i>aipa</i> 'DEM.S.DIST.VIS'	DEM	3b	<i>[ni]-is</i> '3s.F'	deer	Entity	Anaphoric	Identity
6a	<i>?ita</i> 'there.DIST'	PLACE	4a	<i>[ka-]haul</i> 'below'	location deer	Loc/Time	Anaphoric	Antonym
6b	<i>naipa</i> 'DEM.S.DIST.NVIS'	DEM	5a	<i>aipa</i> 'DEM.S.DIST.VIS'	deer	Entity	Anaphoric	Identity
6b	<i>maqmut</i> 'night time'	TIME	4b	<i>[na-]tiŋmuŋ[-in]</i> 'morning'	time of day	Loc/Time	Anaphoric	Metonym

T2: Segment oral narrative

Source: Takivatan Bunun Corpus

Corpus location: TVN-012-001:38-41

Speaker: Tulbus Manququ, M, 75 y

Location and time: Bahuan (Mayuan), 2006

Text

- (14) *Maqai maqabasi tupa tu madaiŋʔaði namuqumaka tanusaŋ matibahi.*
- maqai ma-qabas-i tupa tu ma-daiŋʔað-i*
if DYN-in.former.times-PRT say COMPL STAT-old-PRT
na-mu-quma-ka
IRR-ALL-field-K.DIST
 - tanus-aŋ mati-bahi*
first PROG-have.prophetic.dream
'If in the old days the elders said they wanted to work on the land, they interpreted a prophetic dream beforehand.'
- (15) *Namaqun ʔita maqai masihala bahia, tudip, na, sintupadu tu maqai ʔitun asa na-masihala kakaun.*
- na-maqun ʔita*
IRR-cut.off there.DIST
 - maqai ma-sihal-a bahi-a tudip*
if STAT-good-SUBORD prophetic.dream-SUBORD that.time
 - na sin-tupa-du tu maqai ʔitun*
well RES.OBJ-say-EMO COMPL if there.MED
 - asa na-ma-sihal ka-kaun-un*
be.able IRR-STAT-good CV-eat-UV
'And when they wanted to go there to harvest (lit: when they wanted to cut off things in that place), if the dream was good, that meant in those days that if you were there, you could eat very well.'
- (16) *A maqai dipi madiqla bahia tupa tu asa ni ʔituni nalauq, nitu na ... masihala kakauna sanasia maqai, amin tu maqai ʔitun namuqða kuðaki madiqla bahi, na haiða matað.*
- a maqai dip-i ma-diqla bahi-a*
INTER if then-PRT STAT-bad prophetic.dream-LNK
 - tupa tu asa ni ʔitun-i*
say COMPL have.to NEG there.MED-PRT
 - nalauq ni tu na ma-sihal-a ka-kaun-a*
otherwise NEG COMPL well STAT-good-LNK CV-eat-LNK

- d. *sana-sia* *maqai*
 ACCORDING.TO-ANAPH if
- e. *amin tu* *maqai ?itun* *na-muqða kuða-ki*
 all COMPL if there.MED IRR-again work-K.PROX
- f. *ma-diqla bahi*
 STAT-bad prophetic.dream
- g. *na haiða matað*
 well have die
 ‘And if the dream was bad, then they said that you must not go there, because otherwise you would not eat well, if you followed the rule, but if anyone at all went back to that place to work, and there was a bad dream, people would die.’

(17) *A, maqai mataisaq ... matataisaq a madadaiŋ?að tu, ... maqai mun?ita?a mavia mataisaq tu sadu?uki siatu, sinsusuað bunuað masmamua mavisqai, mavilasa tu-paka madadaiŋ?að tu na maqtu munquma ista?ai nakasihain kakaunun namasi-hala bunun.*

- a. *a* *maqai ma-taisaq*
 INTER if DYN-dream
- b. *ma-ta-taisaq a* *madadaiŋ?að tu*
 DYN-CV-dream INTER elder COMPL
- c. *maqai mun-?ita* *a* *ma-via* *ma-taisaq tu*
 if ALL-there.DIST HESIT DYN-why DYN-dream COMPL
- d. *sadu?u-ki sia* *tu*
 see-K.PROX ANAPH COMPL
- e. *sin-su-suað* *bunuað* *mas-ma-muav* *ma-visqa-i*
 RES.OBJ-CV-SOW plum BE-CV-excessive STAT-abundant.with.fruit-PRT
- f. *mavi-las-a*
 CONTAIN-fruits-LNK
- g. *tupa-ka* *madadaiŋ?að tu*
 tell-K.DIST elder COMPL
- h. *na maqtu* *mun-quma ista-ai*
 well be.possible.to ALL-field 3S.DIST-PRT
- i. *na-ka-sihal-in* *ka-kaun-un*
 IRR-ASSOC.DYN-good-PRV CV-eat-UN
- j. *na-ma-sihal-a* *bunun*
 IRR-STAT-good-LNK people
 ‘And if they dreamt... if the elders dreamt that, if they went over there, they suddenly dreamt that they saw that the plum tree had grown so that it was full of fruits and had large fruits, then the elders would say that it was permitted for them to the land to work, and they would produce good fruits, and the people would also be fine.’

Table 11: Referential cohesion analysis T2

(1)	Reference (Rc)	Rc word class (2)	(3)	Target (Ta)	Referent (4)	Functional role Ta (5)	Phoric function (6)	Rel. Rc-Ta (7)
1a	<i>ma-qabas[-i]</i> 'in the old days'	V					Exophoric	
1a	<i>ma-daip-ʔaδ[-i]</i> 'elders'	N			in former times	Entity	Exophoric	
1a	<i>[na-mun-ʔuma[-ka]]</i> 'go work on the land'	V			land	Loc/Time	Exophoric	
1a	<i>[na-mun-ʔuma[-ka]]</i> 'over there'	ART			location	Loc/Time	Exophoric	
1b	<i>[mati-ʔahi]</i> 'have a prophetic dream'	V			dream	Event	Exophoric	
2a	<i>ʔita</i> 'there.DIST'	PLACE	1a	<i>[na-mun-ʔuma[-ka]]</i> 'go work on the land'	land	Loc/Time	Anaphoric	Identity
2b	<i>bahi[-a]</i> 'prophetic dream'	N	1b	<i>[mati-ʔahi]</i> 'have a prophetic dream'	dream	Event	Anaphoric	Hyponym
2b	<i>tudip</i> 'that time'	TIME	1a	<i>[ma-qabas]-i</i> 'in the old days'	in former times	Loc/Time	Anaphoric	Hyponym
2c	<i>ʔitun</i> 'there.MED'	PLACE			village	Loc/Time	Exophoric	
3a	<i>dip[-i]</i> 'then'	TIME	2b	<i>tudip</i> 'that time'	in former times	Loc/Time	Anaphoric	Metonym
3a	<i>bahi[-a]</i> 'prophetic dream'	N	2b	<i>bahi[-a]</i> 'prophetic dream'	dream	Event	Anaphoric	Hyponym
3b	<i>ʔitun[-i]</i> 'there.MED'	PLACE	2a	<i>ʔita</i> 'there.DIST'	land	Loc/Time	Anaphoric	Identity
3d	<i>[sana-ʔia]</i> 'according to the aforementioned'	V			text	Text	Anaphoric	
3e	<i>ʔitun</i> 'there.MED'	PLACE	3b	<i>ʔitun[-i]</i> 'there.MED'	land	Loc/Time	Anaphoric	Identity
3e	<i>[kuδa]-ki</i> '(the work) in this place'	ART	3e	<i>ʔitun</i> 'there.MED'	land	Loc/Time	Anaphoric	Identity
3f	<i>bahi</i> 'prophetic dream'	N	3a	<i>bahi[-a]</i> 'prophetic dream'	dream	Event	Anaphoric	Identity
4b	<i>madadaipʔaδ</i> 'elders'	N	1a	<i>ma-daip-ʔaδ[-i]</i> 'elders'	elders	Entity	Anaphoric	Identity
4c	<i>[mun-ʔita]</i> '(go) over there'	V	3e	<i>ʔitun</i> 'there.MED'	land	Loc/Time	Anaphoric	Identity
4d	<i>[saduʔu]-ki</i> '(see) here'	ART	3f	<i>bahi</i> 'prophetic dream'	dream	Event	Anaphoric	Metonym
4d	<i>sia</i> 'ANAPH'	ANAPH	4b	<i>madadaipʔaδ</i> 'elders'	elders	Entity	Anaphoric	Identity
4e	<i>bunuaδ</i> 'plum'	N			plum tree	Entity	Exophoric	
4f	<i>[mavi-ʔas[-a]]</i> '(be full of) fruits'	V	4e	<i>bunuaδ</i> 'plum'	plum tree	Entity	Anaphoric	Hyperonym
4g	<i>[ʔupa]-ka</i> '(say) at that time'	ART	3a	<i>dip[-i]</i> 'then'	in former times	Loc/Time	Anaphoric	Part
4g	<i>madadaipʔaδ</i> 'elders'	N	4d	<i>sia</i> 'ANAPH'	elders	Entity	Anaphoric	Identity
4h	<i>[mun-ʔuma]</i> 'go to the field'	V	4c	<i>[mun-ʔita]</i> '(go) over there'	land	Loc/Time	Anaphoric	Identity
4h	<i>ista[-a]</i> '3s.DIST'	PRON	4g	<i>madadaipʔaδ</i> 'elders'	elders	Entity	Anaphoric	Identity
4j	<i>bunun</i> 'people'	N	4h	<i>ista[-a]</i> '3s.DIST'	elders	Entity	Anaphoric	Hyperonym

T3: Segment Biblical narrative

Source: *Tama Dihanin tu Halinga. The Bunun Bible in Today's Taiwan Bunun Version* (Bible Society in Taiwan 2000)

Corpus location: Acts 3:1-10

Text

- (18) *Aiða tu hanian, masa tauḡhuvalin tu ḡintau, Pitilu mas Iuhani hai kusia Masaniṅsiṅ Pisvaṅduan.*

a. *ṽaiða tu hanian*
exist COMPL day

b. *masa tauḡhuvali-in tu ḡin-ta*
WHEN noon-PRV ATTR HOUR-three

c. *Pitilu mas Iuhani hai kusia masaniṅsiṅ pisvaṅduan*
Peter OBL John TPC use holy temple

‘There was a day, when it was three at noon, that Peter and John were using the Holy Temple.’

- (19) *Isia tupaun tu Manauaḡ Ilav ḡia hai, aiða tu taḡini maisna tausṽuvaḡun mapiha, kaupa hanian ansahanun mas bunun mapunsia ilav ḡia, makikisaiv mas nakuṅadah sia Masaniṅsiṅ Pisvaṅduan tu bunun.*

a. *i-sia tupa-un tu manauṽaḡ ṽilav-ḡia hai*
POSS-ANAPH say-UV COMPL beautiful door-DIST.OBL TPC

b. *ṽaiða tu taḡini maisna tausṽuvaḡ-un ma-piha*
exist COMPL one.HUM from give.birth-UV STAT-cripple

c. *kaupa hanian ansahan-un mas bunun ma-pun-sia ṽilav-ḡia*
each day carry.to-UV OBL person DYN-CAUS.ALL-ANAPH door-DIST.OBL

d. *ma-ki-kisaiv mas na-ku-ṽadah sia masaniṅsiṅ pisvaṅduan*
DYN-RED-give OBL IRR-ASSOC.ALL-lower ANAPH holy temple
tu bunun
ATTR person

‘At what was called the Beautiful Gate, there was one man who was cripple from birth, and people carried him every day and put them at that door, and he begged to people that went down into the Holy Temple.’

- (20) *Sadu saia tu Pitilu mas Iuhani hai nakuṅadah sia Masaniṅsiṅ Pisvaṅduan, at maki-saiv naiḡia.*

a. *sadu saia tu*
see 3S.NOM COMPL

- b. *Pitilu mas Iuhani hai na-ku-ḡadah sia masaniḡsiḡ*
 Peter OBL John TPC IRR-ASSOC.ALL-lower ANAPH holy
pisvaḡduan
 temple

- c. *at makisaiv naiḡḡia*
 and make.give 3P.OBJ

‘He saw that Peter and John were about to enter the Holy Temple and made them give (money) [tried to ask them for money].’

- (21) *Naia hai samantuk saiḡḡia tupa Pitilu tu: “Sadua kasu maḡami!”*

- a. *naia hai samantuk saiḡḡia*
 3P.NOM TPC keep.close.watch.on 3S.OBL

- b. *tupa Pitilu tu*
 say Peter COMPL

- c. *sadu-a kasu ma-ḡami*
 see-LNK 2S.NOM DYN-1E.OBL

‘They looked straight at him, and Peter said as follows: “You look at us!”’

- (22) *Saia hai samantuk naiḡḡia, asa usiḡan maḡmaḡ.*

- a. *saia hai samantuk naiḡḡia*
 3S.NOM TPC keep.close.watch.on 3P.OBL

- b. *asa u-siḡa-an maḡmaḡ*
 want ABLE.TO-take-LV which.things

‘He looked straight at them, he wanted to be able to get something from them.’

- (23) *Pitilu hai tupa saiḡḡia tu: “Ukan saikin kim mas sui, haitu nasaivan ku kasu mas inak tu iskakaupa: Mapakasia saikin mas itu takisia Naḡdale tu Iesu Kilistu tu ḡan tupa masu tu, mindaḡkaḡa mudan!”*

- a. *Pitilu hai tupa saiḡḡia tu*
 Peter TPC say 3S.OBL COMPL

- b. *ḡuka-an saikin kim mas sui*
 NEG.have-LV 1S.TOP.AG gold COORD money

- c. *haitu na-saiv-an-ku kasu mas i-nak tu iskakaupa*
 although IRR-give-LV-1S.NSUBJ 2S.NOM OBL POSS-1S.N ATTR everything

- d. *ma-paka-sia saikin mas itu taki-sia Naḡdale tu*
 DYN-RECIP-ANAPH 1S.TOP.AG OBL this.here ORIGIN-ANAPH Nazareth ATTR
Iesu Kilistu tu ḡan
 Jesus Christ ATTR name

- e. *tupa masu tu*
 say 2S.OBL COMPL

- f. *mindan̄kað-a mu-dan*
stand.up-LNK ALL-go
'Peter told him: "I do not have gold or money here, but I will give you everything I have here. I use the name of Jesus Christ who comes from Nazareth to tell you: stand up and walk."'
- (24) *Pitilu hai maʔalak mas isaic̣ɣia tu tanaskaun ima, sidan̄kað saiḍɣia.*
a. *Pitilu hai ma-ʔalak mas isaic̣ɣia tu tanaskaun ʔima*
Peter TPC DYN-lead OBL 3S.POSS ATTR right hand
b. *si-dan̄kað saiḍɣia*
ʔ-stand 3S.OBL
'Peter led him by the right hand, and helped him to stand.'
- (25) *Itu bunun ɣia bantas mas buhtuŋ hai ɣinsu istamasaðan, at mataidaða, matuduldul, kitɲab mudadan.*
a. *itu bunun-ɣia bantas mas buhtuŋ hai ɣinsu*
this.here people-DIST.OBL leg.and.foot OBL joint TPC immediately
is-tamasað-an
TRANSFER-strength-LV
b. *at mataidaða matuduldul kitɲab mu-da-dan*
and jump stand begin ALL-RED-road
'This man here, his legs and joints immediately became powerful, and he jumped up and stood, and he began to walk.'
- (26) *Saia hai taskun naic̣ɣia kuŋadah sia Masaniŋsiŋ Pisvaŋduan, maɣishahainað mudadan, matumashiŋ mas Sasbinað Dihanin.*
a. *saia hai taskun naic̣ɣia ku-ŋadah sia*
3S.NOM TPC do.together 3P.NSUBJ ASSOC.ALL-lower ANAPH holy
masaniŋsiŋ pisvaŋduan
temple
b. *maɣishahainað mu-da-dan matumashiŋ mas Sasbinað Dihanin*
gleeful ALL-RED-road thank OBL God
'Together with them he entered the Holy Temple, and gleefully walk over and he thanked God.'
- (27) *Bunun hai sadu saiḍɣiaa tu mudadan, at matumashiŋ mas Sasbinað Dihanin, at ɣiŋhuða, au pa sahal naia tu saia hai takisia Masaniŋsiŋ Pisvaŋduan tu Manauað Ilav malʔanuhu makisasaiv tu bunun.*
a. *bunun hai sadu saiḍɣia-a tu mu-da-dan*
people TPC see 3S.OBL-LNK COMPL ALL-RED-road
b. *at matumashiŋ mas Sasbinað Dihanin*
and thank OBL God

- c. *at dʒiŋhuða*
and be.startled
- d. *aupa sahal naia tu*
because clearly 3P.NOM COMPL
- e. *saia hai taki-sia masaniŋsiŋ piŋvaŋduan tu*
3S.NOM TPC have.origins.in-ANAPH holy temple ATTR
manau?að ʔilav malʔanuhu sa-makisaiv tu bunun
beautiful door sit.down SEE-beg ATTR person
'People saw him walk, and thank God, and they were startled, because they
recognized him as the man that used to beg sitting down at the Beautiful
Door that was the entrance to the Holy Temple.'

Table 12: Referential cohesion analysis T3

(1)	Reference (Rc)	Rc word class (2)	(3)	Target (1a)	Referent (4)	Functional role Ta (5)	Phoric function (6)	Rel. Rc-Ta (7)
1a	<i>hanian</i> 'day'	N			time of story	Loc/Time	Exophoric	
1b	<i>tauphuvallif-in</i> 'having become noon'	V	1b	<i>dgin-ta</i> 'three o'clock'	time of day	Loc/Time	Cataphoric	Metonym
1b	<i>dgin-ta</i> 'three o'clock'	NUM	1a	<i>hanian</i> 'day'	time of day	Loc/Time	Anaphoric	Part
1c	<i>Pitilu</i> 'Peter'	N			Peter	Entity	Exophoric	
1c	<i>Iuhani</i> 'John'	N			John	Entity	Exophoric	
1c	<i>masanijsin pisvanguan</i> 'the Holy Temple'	N			temple	Entity	Exophoric	Identity
1c	<i>i-sia</i> 'POSS-ANAPH'	ANAPH			temple	Entity	Anaphoric	Part
2a	<i>manau'ad ilavf-dgia</i> 'beautiful door-DIST.OBL'	N	1c	<i>masanijsin pisvanguan</i> 'the Holy Temple'	door	Entity	Anaphoric	Identity
2a	<i>[manau'ad ilavf-dgia]</i> 'beautiful door-DIST.OBL'	ART	2a	<i>i-sia</i> 'POSS-ANAPH'	location temple	Loc/Time	Anaphoric	Metonym
2a	<i>[manau'ad ilavf-dgia]</i> 'beautiful door-DIST.OBL'	ART	2a	<i>i-sia</i> 'POSS-ANAPH'	cripple	Entity	Exophoric	
2b	<i>taqini</i> 'one.HUM'	NUM			cripple	Entity	Exophoric	Identity
2b	<i>mapha</i> 'STAT-cripple'	N	2b	<i>taqini</i> 'one.HUM'	every day	Loc/Time	Exophoric	
2c	<i>hanian</i> 'day'	TIME			people	Entity	Exophoric	
2c	<i>bunun</i> 'person'	N			door	Entity	Anaphoric	Metonym
2c	<i>[ma-pun-jsia]</i> 'DYN-CAUS.ALL-ANAPH'	V	2a	<i>manau'ad ilavf-dgia</i> 'beautiful door-DIST.OBL'	door	Entity	Anaphoric	Metonym
2c	<i>ilavf-dgia</i> 'door-DIST.OBL'	N	2c	<i>[ma-pun-jsia]</i> 'DYN-CAUS.ALL-ANAPH'	door	Entity	Anaphoric	Identity
2c	<i>[ilavf-dgia]</i> 'door-DIST.OBL'	ART	2a	<i>[manau'ad ilavf-dgia]</i>	location temple	Loc/Time	Anaphoric	Metonym
2d	<i>na-ku-nadah</i> 'IRR-ASSOC.ALL-lower'	PLACE	2c	<i>ilavf-dgia</i> 'door-DIST.OBL'	location temple	Loc/Time	Anaphoric	Metonym
2d	<i>sia</i> 'ANAPH'	ANAPH	1c	<i>masanijsin pisvanguan</i> 'the Holy Temple'	temple	Text	Anaphoric	Identity
2d	<i>masanijsin pisvanguan</i> 'the Holy Temple'	N	2a	<i>i-sia</i> 'POSS-ANAPH'	temple	Entity	Anaphoric	Identity
2d	<i>bunun</i> 'person'	N	2c	<i>bunun</i> 'person'	people	Entity	Anaphoric	Cohyponym
3a	<i>sata</i> '3S.NOM'	PRON	2b	<i>mapha</i> 'STAT-cripple'	cripple	Entity	Anaphoric	Identity
3b	<i>Pitilu</i> 'Peter'	N	1c	<i>Pitilu</i> 'Peter'	Peter	Entity	Anaphoric	Identity
3b	<i>Iuhani</i> 'John'	N	1c	<i>Iuhani</i> 'John'	John	Entity	Anaphoric	Identity

(1)	Reference (Rc)	Rc word class (2)	(3)	Target (Ta)	Referent (4)	Functional role Ta (5)	Phoric function (6)	Rel. Rc-Ta (7)
3b	<i>na-ku-nadāh</i> 'IRR-ASSOC.ALL-lower'	V	2d	<i>na-ku-nadāh</i> 'IRR-ASSOC.ALL-lower'	location temple	Loc/Time	Anaphoric	Identity
3b	<i>sia</i> 'ANAPH'	ANAPH	2d	<i>masaniṣṭin piṣvaṇḍuan</i> 'the Holy Temple'	temple	Text	Anaphoric	Identity
3b	<i>masaniṣṭin piṣvaṇḍuan</i> 'the Holy Temple'	N	2d	<i>masaniṣṭin piṣvaṇḍuan</i> 'the Holy Temple'	temple	Entity	Anaphoric	Identity
3c	<i>naidḡia</i> '3P.OBJ'	PRON	3b	<i>Pitilu</i> 'Peter' + <i>Iuhani</i> 'John'	Peter and John	Entity	Anaphoric	Whole
4a	<i>naia</i> '3P.NOM'	PRON	3c	<i>naidḡia</i> '3P.OBJ'	Peter and John	Entity	Anaphoric	Identity
4a	<i>saidḡia</i> '3S.OBL'	PRON	3a	<i>saia</i> '3S.NOM'	cripple	Entity	Anaphoric	Identity
4b	<i>Pitilu</i> 'Peter'	N	3b	<i>Pitilu</i> 'Peter'	Peter	Entity	Anaphoric	Identity
4c	<i>kasu</i> '2S.NOM'	PRON	4a	<i>saidḡia</i> '3S.OBL'	cripple	Entity	Anaphoric	Identity
4c	<i>[ma-]ḡami</i> 'DYN-1E.OBL'	PRON	4a	<i>naia</i> '3P.NOM'	Peter and John	Entity	Anaphoric	Identity
5a	<i>saia</i> '3S.NOM'	PRON	4c	<i>kasu</i> '2S.NOM'	cripple	Entity	Anaphoric	Identity
5a	<i>naidḡia</i> '3P.OBJ'	PRON	4c	<i>[ma-]ḡami</i> 'DYN-1E.OBL'	Peter and John	Entity	Anaphoric	Identity
5b	<i>maḡmaḡ</i> 'which things'	Q	4c		possessions	Entity	Exophoric	Identity
6a	<i>Pitilu</i> 'Peter'	N	4b	<i>Pitilu</i> 'Peter'	Peter	Entity	Anaphoric	Identity
6a	<i>saidḡia</i> '3S.OBL'	PRON	5a	<i>saia</i> '3S.NOM'	cripple	Entity	Anaphoric	Identity
6b	<i>saikin</i> '1S.NOM'	PRON	6a	<i>Pitilu</i> 'Peter'	Peter	Entity	Anaphoric	Identity
6b	<i>kim</i> 'gold'	N	5b	<i>maḡmaḡ</i> 'which things'	possessions	Entity	Anaphoric	Part
6b	<i>sui</i> 'money'	N	5b	<i>maḡmaḡ</i> 'which things'	possessions	Entity	Anaphoric	Part
6c	<i>[na-saiv-an]-ku</i> '1S.N'	PRON	6b	<i>saikin</i> '1S.NOM'	Peter	Entity	Anaphoric	Identity
6c	<i>kasu</i> '2S.NOM'	PRON	6a	<i>saidḡia</i> '3S.OBL'	cripple	Entity	Anaphoric	Identity
6c	<i>i-nak</i> 'POSS-1S.N'	PRON	6c	<i>[na-saiv-an]-ku</i> '1S.N'	Peter	Entity	Anaphoric	Identity
6c	<i>iskakaupa</i> 'everything'	N	6b	<i>sui</i> 'money'	possessions	Entity	Anaphoric	Whole
6d	<i>[ma-paka-]sia</i> 'DYN-RECIP-ANAPH'	V	6b-6c	[entire sentence]	do	Event	Anaphoric	Identity
6d	<i>saikin</i> '1S.NOM'	PRON	6c	<i>i-nak</i> 'POSS-1S.N'	Peter	Entity	Anaphoric	Identity
6d	<i>[taki-]sia</i> 'ORIGIN-ANAPH'	N	6d	<i>Nadale</i> 'Nazareth'	Nazareth	Loc/Time	Cataphoric	Identity
6d	<i>Nadale</i> 'Nazareth'	N	6d	<i>Nadale</i> 'Nazareth'	Nazareth	Loc/Time	Exophoric	Identity
6d	<i>Iesu Kilistu</i> 'Jesus Christ'	N			Jesus	Entity	Exophoric	Identity
6d	<i>ḡan</i> 'name'	N			name	Entity	Exophoric	Identity
6e	<i>masu</i> '2S.OBL'	PRON	6c	<i>kasu</i> '2S.NOM'	cripple	Entity	Anaphoric	Identity
7a	<i>Pitilu</i> 'Peter'	N	6d	<i>saikin</i> '1S.NOM'	Peter	Entity	Anaphoric	Identity
7a	<i>isaidḡia</i> '3S.POSS'	PRON	6e	<i>masu</i> '2S.OBL'	cripple	Entity	Anaphoric	Identity
7a	<i>ḡma</i> 'hand'	N	7a	<i>isaidḡia</i> '3S.POSS'	hand of cripple	Entity	Anaphoric	Part
7b	<i>saidḡia</i> '3S.OBL'	PRON	7a	<i>isaidḡia</i> '3S.POSS'	cripple	Entity	Anaphoric	Identity

1 Referential cohesion in Bunun: A comparison of two genres

(1)	Reference (Rc)	Rc word class (2)	Target (Ta)	Referent (4)	Functional role Ta (5)	Phoric function (6)	Rel. Rc-Ta (7)
8a	<i>bunun[-dɣia]</i> 'people-DIST.OBL'	N	<i>saidɣia</i> '3s.OBL'	cripple	Entity	Anaphoric	Identity
8a	<i>[bunun]-dɣia</i> 'people-DIST.OBL'	ART	<i>na-ku-ɣadah</i> 'IRR-ASSOC.ALL-lower'	location temple	Loc/Time	Anaphoric	Metonym
8a	<i>bantas</i> 'leg'	N	<i>bunun[-dɣia]</i> 'people-DIST.OBL'	limb	Entity	Anaphoric	Part
8a	<i>buhun</i> 'joint'	N	<i>bantas</i> 'leg'	limb	Entity	Anaphoric	Copart
9a	<i>saia</i> '3s.NOM'	PRON	<i>bunun[-dɣia]</i> 'people-DIST.OBL'	cripple	Entity	Anaphoric	Identity
9a	<i>naidɣia</i> '3p.OBJ'	PRON	<i>naidɣia</i> '3p.OBJ'	Peter and John	Entity	Anaphoric	Identity
9a	<i>[ku]-ɣadah</i> 'ASSOC.ALL-lower'	PLACE	<i>bunun[-dɣia]</i> 'people-DIST.OBL'	location temple	Loc/Time	Anaphoric	Metonym
9a	<i>sia</i> 'ANAPH'	ANAPH	<i>masaninsin pisvandung</i> 'the Holy Temple'	temple	Text	Anaphoric	Identity
9b	<i>Sasbinad Dihanin</i> 'God'	N	<i>masaninsin pisvandung</i> 'the Holy Temple'	temple	Entity	Anaphoric	Identity
10a	<i>bunun</i> 'person'	N	<i>masaninsin pisvandung</i> 'the Holy Temple'	God	Entity	Exophoric	Identity
10a	<i>saia</i> '3s.NOM'	PRON	<i>bunun</i> 'person'	people	Entity	Anaphoric	Cohyponym
10b	<i>Sasbinad Dihanin</i> 'God'	N	<i>saia</i> '3s.NOM'	cripple	Entity	Anaphoric	Identity
10d	<i>naia</i> '3p.NOM'	PRON	<i>Sasbinad Dihanin</i> 'God'	God	Entity	Anaphoric	Identity
10e	<i>saia</i> '3s.NOM'	PRON	<i>bunun</i> 'person'	people	Entity	Anaphoric	Identity
10e	<i>[taki]-sia</i> 'ORIGIN-ANAPH'	V	<i>saia</i> '3s.NOM'	cripple	Entity	Anaphoric	Identity
10e	<i>masaninsin pisvandung</i> 'the Holy Temple'	N	<i>saia</i> '3s.NOM'	temple	Entity	Cataphoric	Identity
10e	<i>manauad ?ilav</i> 'beautiful door'	N	<i>masaninsin pisvandung</i> 'the Holy Temple'	temple	Entity	Anaphoric	Identity
10e	<i>bunun</i> 'person'	N	<i>na-ku-ɣadah</i> 'IRR-ASSOC.ALL-lower'	door	Entity	Anaphoric	Identity
10e		N	<i>saia</i> '3s.NOM'	cripple	Entity	Anaphoric	Identity

References

- Abadiano, Helen R. 1995. Cohesion strategies and genre in expository prose: An analysis of the writing of children of ethnolinguistic cultural groups. *Pragmatics* 5(3). 299–324.
- Ariel, Mira. 1991. The function of accessibility in a theory of grammar. *Journal of Pragmatics* 16(5). 443–463.
- Aziz, Yellow Y. 1988. Cohesion in spoken Arabic texts. In Erich H. Steiner & Robert Veltman (eds.), *Pragmatics, discourse and text: Some systemically inspired approaches*, 148–157. London: Pinter.
- Bakhtin, Mikhail M. 1986. The problem of speech genres. In Caryl Emerson & Michael Holquist (eds.). Trans. by Vern W. McGee, *Speech genres and other late essays*, 60–102. Austin, TX: University of Texas Press.
- Benn, Keith Laurence. 1991. *Discourse approaches to cohesion: A study of the structure and unity of a Central Bontoc expository text*. Manila: De La Salle University. (MA thesis).
- Biber, Douglas. 1995. *Dimensions of register variation: A cross-linguistic comparison*. Cambridge: Cambridge University Press.
- Biber, Douglas & Susan Conrad. 2009. *Register, genre, and style*. Cambridge: Cambridge University Press.
- Bible Society in Taiwan (ed.). 2000. *Tama dihanin tu halinga [the Bunun Bible in today's Taiwan Bunun Version]*. Taipei: The Bible Society in the R.O.C.
- Connor, Ulla. 1984. A study of cohesion and coherence in English as a second language students' writing. *Paper in Linguistics* 17(3). 301–316.
- Crossley, Scott A. & Danielle S. McNamara. 2012. Predicting second language writing proficiency: The roles of cohesion and linguistic sophistication. *Journal of Research in Reading* 32(2). 115–135.
- De Busser, Rik. 2009. *Towards a grammar of Takivatan Bunun: Selected topics*. Melbourne: La Trobe University. (Doctoral dissertation).
- De Busser, Rik. 2013. The influence of Christianity on the Bunun language: A preliminary overview. In *Proceedings of the International Workshop on "Special Genres" in and around Indonesia*, 59–76. Tokyo: Research Institute for Languages, Cultures of Asia & Africa, Tokyo University of Foreign Studies.
- De Busser, Rik. 2017. Spatial deixis, textual cohesion, and functional differentiation in Takivatan Bunun. *Oceanic Linguistics* 56(1). 89–121.
- Ezard, Bryan. 1978. Insights on cohesion from Tawala. *Oceanic Linguistics* 17(2). 107–132.
- Flaming, Rachel. 1983. Cohesion in Wandamen narrative. *NUSA: Linguistic Studies of Indonesian and Other Languages in Indonesia* 15. 41–49.
- Foley, William A. 2007. The place of Philippine languages in a typology of voice systems. In Peter Austin & Simon Musgrave (eds.), *Voice and grammatical relations in Austronesian languages*, 22–44. Stanford, CA: CSLI Publications.
- Fox, Barbara A. 1987. *Discourse Structure and Anaphora: Written and Conversational English*. Cambridge: Cambridge University Press.
- French, Koleen Matsuda. 1987. The focus system in Philippine languages: An historical overview. *Philippine Journal of Linguistics* 18(2)/19(1). 1–27.

- Givón, Talmy. 1993. Coherence in text, coherence in mind. *Pragmatics and Cognition* 1(3). 171–227.
- Haiman, John. 1983. Iconic and Economic Motivation. *Language* 59. 781–819.
- Halliday, M. A. K. 1994. *An introduction to Functional Grammar*. 2nd edn. London: Arnold.
- Halliday, M. A. K. & Ruqaiya Hasan. 1976. *Cohesion in English*. London: Longman.
- Halliday, M. A. K. & Christian M. I. M. Matthiessen. 2004. *An introduction to Functional Grammar*. 3rd edn. London: Hodder Arnold.
- Halliday, M. A. K. & Christian M. I. M. Matthiessen. 2014. *Halliday's introduction to Functional Grammar*. 4th edn. London: Routledge.
- Hassel, R. Chris. 2005. *Shakespeare's religious language: A dictionary*. New York: Continuum.
- Hasselgård, Hilde. 2004. The role of multiple themes in cohesion. In Karin Aijmer & Anna-Brita Stenström (eds.), *Discourse Patterns in spoken and written corpora*, 65–87. Amsterdam: John Benjamins Publishing Company.
- Hickmann, Maya & Henriëtte Hendriks. 1999. Cohesion and anaphora in children's narratives: A comparison of English, French, German, and Mandarin Chinese. *Journal of Child Language* 26(2). 419–452.
- Himmelman, Nikolaus P. 1996. Demonstratives in narrative discourse: A taxonomy of universal uses. In Barbara Fox (ed.), *Studies in anaphora*, 205–254. Amsterdam: John Benjamins Publishing Company.
- Hoey, Michael. 1991. *Patterns of lexis in text*. Oxford: Oxford University Press.
- Huang, Hui-chuan J. (黃慧娟) & Chao-kai (施朝凱) Shih. 2016. *An Introduction to Bunun Grammar*. Taipei: Council of Indigenous Peoples.
- Huang, Yan. 2000. *Anaphora: A cross-linguistic study*. Oxford: Oxford University Press.
- Kehler, Andrew. 2002. *Coherence, reference, and the theory of grammar*. Stanford, CA: CSLI Publications.
- Kehler, Andrew. 2004. Discourse coherence. In Laurence R. Horn & Gregory L. Ward (eds.), *The handbook of pragmatics*, 241–265. Malden, MA: Blackwell.
- Kruger, Alet. 2000. *Lexical cohesion and register variation in transition: "The merchants of Venice" in Afrikaans*. Pretoria: University of South Africa. (Doctoral dissertation).
- Lambrecht, Knud. 1994. *Information structure and sentence form: Topic, focus and the mental representations of discourse referents*. Cambridge, UK: Cambridge University Press.
- Li, Paul Jen-Kuei. 1988. A comparative study of Bunun dialects. *Bulletin of the Institute of History and Philology, Academia Sinica* 59(2). 479–508.
- Li, Paul Jen-Kuei. 2008. The great diversity of Formosan languages. *Language and Linguistics* 9(3). 523–546.
- Malah, Helen Tan, Zubairu & Sabariah Md Rashid. 2016. Evaluating lexical cohesion in Nigerian newspaper genres: Focus on the editorials. *International Journal of Applied Linguistics and English Literature* 6(1). 240–256.
- Mann, William C. & Sandra A. Thompson. 1987. *Rhetorical structure theory: A theory of text organization*. Marina del Rey, CA: Information Sciences Institute.

- Martin, James R. 1992. *English text: system and structure*. Philadelphia: John Benjamins Publishing.
- Martin, James R. 2001. Cohesion and texture. In Deborah Schiffrin, Deborah Tannen & Heidi E. Hamilton (eds.), *The Handbook of discourse analysis*, 35–53. Malden, MA: Blackwell.
- Merrell, Floyd. 2001. Charles Sanders Peirce's concept of the sign. In Paul Cobley (ed.), *The Routledge companion to semiotics and linguistics*, 28–39. London: Routledge.
- Quine, Willard V. 1985. Events and Reification. In E. Lepore & B. McLaughlin (eds.), *Actions and events: Perspectives on the philosophy of Davidson*, 162–171. Malden, MA: Blackwell.
- Riesberg, Sonja. 2014. *Symmetrical voice and linking in Western Austronesian languages*. Berlin: Mouton de Gruyter.
- Tanskanen, Sanna-Kaisa. 2006. *Collaborating towards coherence: lexical cohesion in English discourse*. Amsterdam: John Benjamins Publishing.
- Zacks, Jeffrey M. & Barbara Tversky. 2001. Event structure in perception and conception. *Psychological Bulletin* 127(1). 3–21.
- Zeitoun, Elizabeth. 2000. *A reference grammar of Bunun* (布農語參考語法). Taipei: 遠流.